



Left Inverted Terminal Repeat: 1-103

Encapsidation Signal (Ψ): 183-331

HPRT Introns: 365-10557

Right Inverted Terminal Repeat:
10571-10673

pBR322 ori: 10877-11544

Kanamycin Resistance Gene: 12353-
13144

FIG. 1

SEQ ID NO:1 pSHuttle Sequence

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGGGGGTGGAGTTTGTGACG
TGGCGCGGGGCGTGGAACGGGGCGGGTGACGTAGTAGTGTGGCGGAAGTGTGATGTTGCAAGTGTG
GCGGAACACATGTAAGCGACGGATGTGGCAAAAGTGACGTTTTTGGTGTGCGCCGGTGTACACAGGA
AGTGACAATTTTCGCGCGGTTTTAGGCGGATGTTGTAGTAAATTTGGGCGTAACCGAGTAAGATTTGG
CCATTTTCGCGGGA AAACTGAATAAGAGGAAGTGAAATCTGAATAATTTTGTGTTACTCATAGCGCGT
AATACTGGTACCGCGGCCGCTCGAGTCTAGAAGTGTGGATCCCCCGGGCTGCAGGAATTCTGATGG
CTCTCAAAATTCCTGCCTCCTTTAGGGATAAAAAGACTTTAAGACTTTTTAACAAAAAAGAAAAAGAAA
AAAAAAATTCCTGCCTCCTGGTGTACACACACAGAAGGGTTCCCTCCCCTTGAATGTGACCAGGATCT
GTGAAAATAACGGGATAGCCGCTCCTGTGATTAGGTTATGTGGTAGACTAGAGCAAGATTCTCCTGCT
GGTTTTGAAGAAGTCAGCTGCCATGTTGTGAGACTGTCATGGGCTAGGGCATGAGCCTTTAAATATCT
GGGAGCAACCCCTGGCCAGCAGCCAGTGAGAAAACGGGGCCCTCAGTCCTACAATCACAAGGAATAA
ATTCTGCCAACAACTGAAGGAACCTTTGAAGAGGATCATGAGTCCCTTGATTCAGCTTGATGAGCCCC
TGAGCAGAGGATACAGCTAACTTGTACTAGGGAAGTATAAAAAACATGCATGGGAATGATATATATC
AACTTTAAGGATAATTGTCATACTTCTGGGAATGAAGGGAAAGAAATGGGGCTTTAGTTGTATTATGA
TCTTTAATTTCTCAAAAAAATAAGATCAGAAGCAAATATGGCAAAATGTTAATACTTTTGTGGGTAC
GTAGGTATTGAGCATACCTTTTTTTCTGAGTTCAAAATATTTTATAATTAAATGAAATGCAGGCCAGG
CACAGTGGCTCATGCCTATAATACCAGCACTTTGCGAGGCCGAGGTGGGAGGATGGCTTGAGGCCAGA
CCAGCCTGGCCAACATGGCAAAACCCCATCTCTACTTAAAAAAAAAAAACTATATATATATATATGT
GTGTGTGTGTGTATATATATATATGTATATATATTTATATATGTGTGTATATATATATATGTATATAT
TTATATATGTGTGTGTATATATATATATATACACACACACACATATATACATACATACACACACACA
CACACACAATTAGCCAGGCATGGTGGCGCACACCTGTAGTCCCAGCTACTTGGGAGGCTGAGACATGA
GAATTGCTTGAACCTGGGAGGCAGAGTAGTTAGTGAGCTGAGATCATACCACTGCACTCCAGCCTGGT
GACAGAGTGAGACTCTGTCTTAAAAAAAAATAAAATTTAAATTTAAATGCAAAAGGTCCAAGTGAATT
GAAGAGGAAAGGGGTATCAAGGAAGGTTTTGTGGAGGTGACGTTTGAGCTGGGTCTTAAATGACTTA
AACATGGGATAAGAAGGGAGGGAATAAGGACATTTGAGGTACGAGAAATAAGGAGCATCAGTGGA
ACAACCTAACGTCTGTCAACCAAGTGAATGGATAACAAAATGTAATTCAGATGGTATCCAACCTACGA
TGGTTCCAACATGAGATTTTTCTGACTTTAGGATAGATTTATCAAAGTAGTAAATCCATTTTCAACTTA
TGATATTTTCAACTTCAGATGGGTTTATCAGGACACAGTTGAGGAACACCTGTCTATCCATACAATTTG
GCAATAAAAAGGAAATGAGTGCAGATATACTCCACAACATGAATGAACCTTGAAAACATTAAGTGAG
AGAAGCCAGATACAAAAGGCCACATATTGTATGATTCTATTTATACAAAATGTCCAGAATAGGCAAAT
CTTATAGACAGCAAGTAGGTAGATGATCAGTTTGCTAGGTGCTGGGGGAAGGGGAAATGGGGAGTGA
TGGCTAAGGGGATTGGGTTTTCTTTGTGGGGAAATGAAAATGTTTTAAATTTGAGCGTGATAATGATTG
CTCAATGCTGCATATATATATAATCTATAGATTATATATATATAAAGAGAGGCTGTTAGACAGTGATA
AGTGATATATATATATATATACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGCTGTTAGTGATAAGTG
ATCAGGAAAATAAAAAGTATTGAGGAGGAATACGAAGTTGACGGTGTGAAAACATGAGATTTTATATA

GGATGGCCAGGGAAGGCCTTAATGAGAAAGTGACTTATGAGTAAAAACAAGGGATCCTAAACCTTAG
CATGCATCAGAATCACTCGGAAACTTGTTAAAGCATAGCTTGCTGGGCCTCATCACAGATATTTTGATT
CGGTAGGTTCTTGTCTGATATTAATACTTTTGGTCTAGGGAACCACATTTTGAGAACCACTGAGCTAAA
GGAAGTAAAGGTTTCCCTTAGTTTACTAGCTGGTAACCCTAGGAAACTGCTTAGCCTCTCGGTGCTAAG
ATACAAAATACTTTAGCACATAATAACACATGGAAAATAGTCTATAAATTATAAATATTATTTTTTATG
TACCAAATATTACATAAGACAAAATCTAAGCAAGATATATATATATATACATAAAATATAAGATATAT
ATGTATATATTATATATAGATAAATAGAGAGAGAGAGTTATGTTTAGAAAGAAAATACTTCAAACCTAA
AAAAAGAGAGGTAGGAAGTATACCATTCCATTATTGGTAAAAACAAATTACTAAGTAGTCTTTACAAA
AAACCAATCTCACTCCTTTAGAACACAAGCCCACCATTAAAACCTGATGCAGAGGAATTTCTCTCCCTG
GCTTACCTTTAGGATGGTGCATACTAAGTTAGAAAAGTCATAAATGTTATATTAAAAGTAAATGTGAA
CTTACTTCCACAATCAAGACATTCTAGAAGAAAAAGAGAAATGAAAATCAGTACAATGAATAAAACG
GTATTTCCAATTATAAGTCAAATCACATCATAACAACCCTAAGGAATTATCCAAACTCTTGTTTTTAGA
TGCTTTATTATATCAAACCTCTCCTTTAAACAAGTGGCCCATCTGCTGGGATTTGGAAGCCTGTAATACT
GAAATTTTCATCATAATGGAAATTTTAAAAACAGAATTTGACCCACCTGTTTTTAAAAACACTTTTCATTA
CTTAACAAGAGGTCTAATCTTGGGCAAGTCTTGAAATTTCTCTGGCCTTAGTTTCCCATGTGTAAATG
AACTTGAAGCAGTTGGTCTCTTATAGTCTCCTGACTCTAACATTCTAAGAATTATATTTGTACAATAA
CTCAAAAATCACATAATTTAATTTACCATATGGACTCCAAAATATATTTTCTCATTAGGCTAAACTTGA
TCTGCATTTTCTGGATGTGTCCATATTCTTGGACTACACTAAAACATGATACCAATGCTTCCTCTCACC
ATAAACCTCCTCACTTCGCTTTCTACATTTAAGAATTTTATAGCTGGAAGAGTCCTTAACAGAAAATACCA
TCTAATAATTACCCCTCAAAATCGAGAAAGTCCTATCTGTTCTTATGCTAGTTATAAGAATGAGGCAGC
ATTTACATAATGGTTATAAACACTGCCACAAGAAGATTCATGATGTGTTGTTTATCTGTAGCTCTCAT
CATACTCTGTCATATAACTATAGCATTAAAGATTTTAATGTTCTATATATTCTTCTAAGACAGTGTTTACC
AGAGTAAGGCACAAAAGATCCACTGGTTTGCAAGAAAGATTAGAACTTTTAAATTTTTTAACTCACC
TTGTTTAATCTATATTTTTGTATGTATTTTGTAAACATATATATTATTATTACCATAAATCATATATAATTT
AAAATGCATATATTAGGGGTAAATGCTCAGGAAACTTTTTATAAATTGGGCATGCAAATACAAGTTTG
AAGACTCACTGTTCTAGGTATTAAGTAAAGTTATAACCAAGTAAAGCTTCCACCTTTTCATGTCTCA
AAGCAGTTTATTGTTGGAGGTAAGATCTCTTAGAAGCCTAAACAGGTCCAAGTACAGAATGAAGTAAG
GCTAGCCCATAACTTGTGGCAAGCAATTCATACTATTTCTCTCATGCTGAGCTCTCCTCAGTGAAGCAG
CTACTATAGACAACTGCAGCCTATTGGTAGCCTATTTTACAGGCAGGAAAAAAATTACTTTTTTATTCA
AAGTGGAACCTCAGGACATGGGGAGAAAATGAATACAAAAAATAGGGTCAATCCAAAGGCACACAGC
AAATGAGTAACACAGTTATGTTTTTTTTCCCATTTGTATGAGGTCCAGTAAATTCTAAGTAAACTGCAA
ATTTAATAATACACTAAAAAAGCCATGCAATTGTTCAAATGAATCCCAGCATGGTACAAGGAGTACAG
ACACTAGAGTCTAAAAAACAAAAGAATGCCATTATTGAGTTTTTTGAATTATATCAAGTAGTTACATCT
CTACTTAATAAATGAGAAAAACGAGGATAAGAGGCCATTTGATAAAATGAAAATAGCCAAGAAGTG
TATTAGAGACTTGAATACAGGTATTCGGGTCCAAAGTTCATCTGCTCAAATACTAACTGGGGAAAAGA
GGGAAAAATATTTATATACATATATATCTGCACACAAAAATACCCCCAAAAGACAAAATGAGGCCAG
GCAGGGTGGCTCACACCCGTAATCCCGGTACTTTGGGAGGCTGAGGCAGGTGGATACCTGAGATCAGG

FIG. 2A

AGTTGGAGATCAGCCTGGTCAACATGGTGAACCCTGTCTCTACTAAAGATAAAAAAATTAGCCAGGC
ATGGTGGCGTGCGCCTGTAATCCCAGCTACTTGGGAGTCTGAGGCAGGAGAATCACTTGAAGTGGGAA
GGGGAGGTTGCAGTGAGCCAAGATCGTACTACTGCACTCCAGCCTGGGCAGCAGAGTGAGACTCCATC
ACAAAAATAAATAAATAAATAAATAACAATGAAACAGAAAGTTCAAATAATCCCATAATCTTACCAC
CAAGAAATAACTTTCACTCGTTATACTTATTGATTTTTCCATAATAAATGTACTTTACTGTGACTATCAT
GAAAAGAAAGTTATTTTAGAAACAGAGAACTGTTTCAGATCAAATCTATGTAGTAGAACAGAGCCATT
AGGTGGGAAAGACGAGATCAAATAATCTCAGAAGGCCTAAAAGGCTAGGTCCATTCCAGCACTAA
AAACTGACCAGACAAGTAATGGCTTCAACAGCTTCTAAATATGGACAAAGCATGCTGAAAGGGAAGG
ACAGGTCTAACAGTGGTATATGAAATGAACAGGAGGGGCAAAGCTCATTCTCTCTGAAGTTTTCCA
AAGATGCTGAGGAGGACATTAGTTTGACATGACCCTGATATGGGACAAGATAATTCACAGAAGTTTT
ACATGTTAAAGTTTTCTTATAGATACTCATTCAAGTAAGCAATGAACACTAAAATCTAAAGAAAGAAA
AGAGCTTTAGAGTCAGGTCTGTATTCAAATTCAGCTCTACCACTTACTGGTTCTGTGACTTTGGGCAA
GTCTTTTAACCTTATTAAGTCTTAATTTCTGATTTGTAAAATGGGGATATCGTCTCCCTCACAGGATTG
TTGTGAAACTTTTATGAGATTAATGCCTTTATATTTGGCATAGTGTAAGTAAACAATAACTGGCAGCTT
CAAAAAAAAAAAGCAGTAGCATTCCATCATTATTATTGGTTACTCTCAAAAAGTTTTTCAATGTACTA
GAAGATAAATATTCAAATACCTTAATATCTCCATTATTTTCAGGTAAACAGCATGCTCCTGAACAACCA
ATGGGTCAACAAATAAATTAAGGGAAATCTAAAAACATCTTGATATTAACTACATGGAAGCACA
ATATACCAAACCAATGGTTCACACTAGGAGAATTTTAAGGTACAAGAAAACCTCTTTGAGATTTCTTA
AAATAATAGTATGTCTGAATTTATTGAGTGATTTACCAGAACTGTTGTAAGAGCTCTACTTGCATTAT
AGCACTTAATCCTCTTAACCTCTATGGCTGCTATTATCAACCTCACCTAATCACATATGGGACACAGAG
AGGTTAAGTAACTTGCCCAAGGTCAGAGTTAGGAAGTACTAAGCCATGCTTTGAATCAGTTGTCAGGC
TCCGGAACCTCACACTTTTCAGCCACTACATAATACTGCTTTGCTATCTTTTAGGAAACTATGTGAGTCTA
CCTCACATAGACTCACATAGGTTTGTTTTTTTTTTTTTTTTAAAGGCTATCTTTTCCCCCATCAATGTTTT
TTGAAGGATCCCAAATTAGAGTCCACAGAGGCAGACAGCAGTACTTGACAATATGGACATTTAAGGT
TAATGTTGGATTCTACTGTCTTTTTACTACATGACCTAGGGAACGATAATTAACCTAGACTGCTTCCAA
GGGTAAATAACCCATTTAGTTATACTATGTAAATTATCTCTTAGTGATTGATTGAAAGCACACTGTTA
CTAATTGACTCGGTATGAAGTGCTTTTTTTTCTTCCCTTTCAAGATACATACCTTTCCAGTTAAAGTTGA
GAGATCATCTCCACCAATTACTTTTATGTCCCTGTTGACTGGTCATTCTAGTTAAAAAAAAAAAAAACT
ATATATATATATATCTACACACACATATGTATATGTATATCCTTATGTACACACACAACTTCAAATTA
AATGAGAACTAGAAGATTTGAGAAGTTAGCTAGCTAATATCCATAGCATTATGATATTCTAAATGATA
TGAATTATAAGAATTAGGTTTCCTGAAATGAATGACTAGAAAACCTTCAAGTAGAGATTAGTAAAAAT
TAAAAAGTCCTAATCGGCCATTACTGATTTGATGTTTTTAAGAGTCCTAAAAAATGGGTACATCCATT
TTTAAGTGGGTAGTATTATAACAGCCACCCATCTTCAATCACAGTGATTTCTGAATTGTGAGGGGAAGTT
ATTAGCATGACAGGTGTCTGGTTCTGGCCCTGTACGATTCCTATGAGTCAAGCAAATTGTAAGGGCTG
GTCTATATCACACCCAACCCCAAGGATATGTCCCTCAAAGTCTAGCCCAGGCCCGTCATCTTCAGC
ATCATCTGGGAAACCAGGTCTGATTAGTAGTCCTTTAAGGAATACCTCTTAGGCTCCCATTTTACTGCT
ATCACAGAATCCAATAAAACCCTTACAGGAGATTCAATGGGAAATGCTCAACACCCACTGTAGTTGGT

FIG. 2B

GGTGACAATGACCATAATTTGGCTGTGCTGGATTTCAGGACAGAAAATTTGGGTGAAAGAGCAGGTGA
ACAAAAGAGCTTCGACTTGCCCTAGCAGAGAGCAAGCCATACCATACCACAAAGCCACAGCAATTAC
AACGGTGCAGTACCAGCACAGTAAATGAACAAAGTAGAGCCCAGAAACAGACCCAGAACTATATGAG
GATTTAGTATACAATAAAGATGGTATTTTCGAGTCAGTAGGGAAAAGATGAATTATTCAATAAATGATG
TTTGGCCAAC TAGTAACCCATTTGGGAAAAAATAAAAGTATGGTCCCTACCTCACAGCATAACAAAA
ATAAATTCCAGACGGATTAAATCTAAATGTAAAAAATAAAGCCATAAGTGGACTGGAAGAAAATAG
AGAATTTTTTTTAACATCCGTAGAAAGGGTAAAAACCCAGGCATGACATGAACCAAACTGAAGAGG
TTCTGTAACAAATACCCCTTTTATATATTGGGCTCCAACAATAAGAACCCATAGGAAAATGGAGAAT
GAACACAAATAGACAATTTATAGAAGAGAAGGTTATAAGGTGTAAATTTATATCTATCTGAGAAACA
AACACTAAAACAATGTGATTCTACTGTTCTCCACCCATACTGGCAAACTTAAGCCTGATAATATGCT
GAGGGGAAATAAGCACTCTTGTTGGTGAGAGTATTAATTGGCATAGCTTCTTTTGAAAATGACATAGC
AATACCTGTTAAAATTGCAAACATGCATGTCACTTAATTCCATGTAATTCCTACTTCTGGGAATCAATT
GCTACAAAAACACTTGACAAGTATACAAAGATACATTCAAGAGTGTTCACTGGGCCGGGTGCGGTGGC
TTCATGCCTGTAATCCCAGGGAGGCAGAGGCAAGACGATCGCTTGACCCAGGAGTTCAAGGCCAGCC
CGAGAAACACAGCAAGACCCTGTCTCTCTTTTTTTTATTTAAAAAATAAATGTTCACTGTATCAGTTGT
TCACAAAAACAAACCAACATGTCCATTAACAGGGAACCATTTAAATTAATCAAGTTCATCTACACAAT
GTAATACCATGCAACTATTA AAAAGCACCTGATAATCCAAAGCACACTGAGACAGAATAATGCTATTA
AAAACACCAAGTAGTGGAACACTGTGTTGCCTATGACACCATTTTTATTCAACATTTAAACAAATTTGT
AACAGCAATTACATGAGTAGTGACAATGGCGTTTATGAGACTTTTCACTTTTATGTGCTTCTATTTTTGT
TATGCTTCTATATATACATCCATTTATTATGGAGTGTTACTTTCAAAAATCACAATGGGCCAGTATTA
TTTGGTGTTGCAAGGTGAGCATATGACTTCTGATATCAACCTTTGCATATTACTTCTCAATTTAGGGAA
ATTACAGACATCCCTTATTCTAACTAACTTAAAACCCAGCATTTCAAACATACAGAATTGATGGGGAA
AAAAAAGAAAGAAGAAAGAAAGAAAAGGCAACAAGCTTCAGATGACAGTGACTCACATCAAATTATT
TATAAAATCTGTAAATAGTGCCATCTTCTGGAGATACCTGGTATTACAGTCCAACTCCAGTTGATGTC
TTTACAGAGACAAGAGGAATAAAGGAAAAAATATTCAAGAAGTGAAGATGAGGAGTCATGGAAAAA
TTGCTGTGATCCAAAGGCTACGGTGATAGGACAAGAAACAAGAGAACTCCAAGCAGTAAGACACTGC
TGTTCTATTAGCATCCAAACCTCCATACCTCCTGTTTGCCCCAAGGCTTTTTTAAAAAATAGAGACAGG
ATCTCACTATTTTGCTCAGGCTGGTCTTGAACCTCCTGGACTCAAGCTATCCTCCTGCCTCGGCCTCCTAA
AGTGCCGAGATTACAGGCTTGAGTCACCATACCTGGCTATTTATTTTTCTTAACTCTCTTGCCTGGCCT
ATAGCCACCATGGAAGCTAATAAAGAATATTAATTTAAGAGTAATGGTATAGTTCACTACATTGGAAT
ACAGGTATAAGTGCCTACATTGTACATGAATGGCATACATGGATCAATTACCCACCTGGGTGGCCAA
AGGAACTGCGCGAACCTCCCTCCTTGGCTGTCTGGAACAAGCTTCCCACTAGATCCCTTTACTGAGTGC
CTCCCTCATCTTTAATTATGGTTAAGTCTAGGATAACAGGACTGGCAAAGGTGAGGGGAAAGCTTCCT
CCAGAGTTGCTCTACCTCTCCTCTACCGTCCTATCTCCTCACTCCTCTCAGCCAAGGAGTCCAATCTGT
CCTGAACTCAGAGCGTCACTGTCAACTACATCAAAATTGCCAGAGAAGCTCTTTGGGACTACAAACAC
ATACCCTTAATGTCTTTATTTCTATTTTGTCTACCTCTTCAGTCTAGGTGAAAAAATAGGAAGGATAAT
AGGGAAGAACTTTGTTTATGCCTACTTATCCGCCCCTAGGAATTTTGAAAACCTCTAGGTAGCAATAA

FIG. 2C

GAAGTGCAGCATGGTATAGAAAAAGAGGAGGAAAGCTGTATAGAAATGCATAATAAATGGGCAGGA
AAAGAACTGCTTGGAACAAACAGGGAGGTTGAACTATAAGGAGAGAAAGCAGAGAGGGCTAATCAAC
AAGGCTGGGTTCCTAAGAGGGCATGATGAGACTATTACTAAGGTAGGAATTACTAAGGGCTCCATGTC
CCCTTAGTGGCTTAGTACTATGTAGCTTGCTTTCTGCAGTGAACCTCAGACCCTTCTTTTAGGATCCTAG
AATGGACTTTTTTTTTTATCGGAAAAACAGTCATTCTCTCAACATTCAAGCAGGGCCCCAAGTCTACCAC
ACTCAATCACATTTTCTCTTCATATCATAATCTCTCAACCATTCTCTGTCCTTTTAACTGTTTTTCTATAC
CCTGATCAAAATGCCAACAAAAAGTGAGAATGTTAGAATCATGTATTTTTAGAGGTAGACTGTATCTCA
GATAAAAAAAAGGGGCAGATATTCCATTTTCCAAAATATGTATGCAGAAAAAATAAGTATGAAAGG
ACATATGCTCAGGTAACAAGTTAATTTGTTTACTTGATTTTTATGAATTCCCTAAAACCTACGTCACCC
GCCCCGTTCCACGCCCCGCGCCACGTCACAACTCCACCCCCTCATTATCATATTGGCTTCAATCCAA
AATAAGGTATATTATTGATGATGTTAATTAACATGCATGGATCCATATGCGGTGTGAAATACCGCACA
GATGCGTAAGGAGAAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCG
GTCGTTCCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGG
GGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCG
GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA
GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
CCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGCGCTTTCTC
ATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCGTTACAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG
ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA
GAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCT
GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCG
GTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC
TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATC
AAAAAGGATCTTCACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATG
AGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTC
GTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCC
CCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCA
GCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTG
CCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTGCAGCCA
TGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTACGTAGAAAGCCAGTCCGCAGAAACGGTGC
TGACCCCGGATGAATGTCAGCTACTGGGCTATCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAGC
AGGTAGCTTGCAGTGGGCTTACATGGCGATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCG
GAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAACTGGATGGCTTTCTT
GCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTCGC
ATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGA
CTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCCG

FIG. 2D

TTCTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGCTATCG
TGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTG
GCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTAT
CCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTGACCACCAA
GCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGA
CGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTGCGCCAGGCTCAAGGCGAGCATGCCCCGACGGC
GAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGA AAAATGGCCGCTTTTCT
GGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGA
TATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGA
TTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCCTCTGAATTTTGTTAAAATTTTGTAAAT
CAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGA
TAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAA
GGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCTAATCAAGTTTTTTGGG
GTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGA
AAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCA
AGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTC
CATTCGCCATTCAGGATCGAATTAATTCTTAATTAA

FIG. 2E

SEQ ID NO:2 Human TM amino acid sequence

MLGVLVLGALALAGLGFPAPAEPPQPGGSQCVEHDCFALYPGPAT
FLNASQICDGLRGHLMTVRSSVAADVISLLLNGDGGVGRRRLWIGLQLPPGCGDPKR
LGPLRGFQWVTGDNNTSYSRWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQQ
CEVKADGFLCEFHFPAATCRPLAVEPGAAAAAVSITYGTPFAARGADFQALPVGSSAA
VAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGCEHACNAIPGAPRCQCPA
GAALQADGRSCTASATQSCNDLCEHFCVPNPDQPGSYSCMCETGYRLAADQHRCE
VDDCILEPSPCPQRCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYQCQPLN
QTSYLCVCAEGFAPIPEPHRCQMFCNQTACPADCDPNTQASCECPEGYILDDGFICT
DIDECENGGFCSGVCHNLPGTFCICGPDSALARHIGTDCDSGKVDGGDSGSGEPPPS
PTPGSTLTPPAVGLVHSGLLIGISIASLCLVALLALLCHLRKKQGAARAKMEYKCAA
PSKEVVLQHV RTERTPQRL

FIG. 3

SEQ ID NO:3 human TM nucleotide sequence

atgcttggg gtcctggcc ttggcgcgt ggccctggcc ggcctgggt tccccgcacc cgcagagccg cagccgggtg
gcagccagtg cgtcgagcac gactgcttcg cgctctaccc gggccccgcg accttctca atgccagtca gatctgcgac
ggactgcggg gccaccta at gacagtgcgc tctcgggtgg ctgccgatgt catttcttg ctactgaacg gcgacggcgg
cgttggccgc cggcgcctct ggatcggcct gcagctgcc cccggctgcg gcgaccccaa gcgcctcggg cccctgcgcg
gcttccagtg ggttacggga gacaacaaca ccagctatag cagggtgggca cggctcgacc tcaatggggc tccccctgc
ggcccggtgt gcgtcgctgt ctccgtgct gaggccactg tgcccagcga gccgatctgg gaggagcagc agtgcgaagt
gaaggccgat ggcttctct gcgagttcca ctccagcc acctgcaggc cactggctgt ggagcccggc gccgcggctg
ccgccgtct gatcacctac ggcaccccgt tcgcggcccg cggagcggac ttccaggcgc tgccgggtggg cagctccgc
gcgggtggctc cctcggctt acagcta atg tgcaccgcgc cggccggagc ggtccagggg cactgggcca gggaggcgc
gggcgcttg gactgcagc tggaacacgg cggctgcgag cagcgtgca atgcgatcc tggggctccc cgctgccagt
gcccagccgg cggcccttg caggcagacg ggcgtcctg caccgcatcc gcgacgcagt cctgcaacga cctctgcgag
cactctgcg tccaaccc cgaccagccg ggctcctact cgtgcatgtg cgagaccggc taccggctgg cggccgacca
acaccggtgc gaggacgtgg atgactgcat actggagccc agtccgtgc cgcagcgtg tgtcaacaca cagggtgggt
tcgagtcca ctgctaccct aactacgacc tggtgacgg cagtggtgtg gagccgtgg acccgtgctt cagagccaac
tgcgagtacc agtgccagcc cctgaacaa actagctacc tctgcgtctg cggcagggc ttgcgcccc tccccacga
gccgcacagg tgccagatgt ttgcaacca gactgcctgt ccagccgact gcgacccaa caccaggt agctgtgagt
gcccgaagg ctacatctg gacgacggt tcatctgcac ggacatgac gagtgcgaaa acggcggctt ctgctccggg
gtgtgccaca acctccccg taccttcgag tgcattgcg ggcccgaact ggcccttgcc cgccacattg gcaccgactg
tgactccggc aaggtggacg gtggcgacag cggctctggc gagccccgc ccagcccgc gcccggctcc acctgactc
ctccggccgt ggggctcgtg cattcgggt tgcctatagg catctccatc gcgagcctgt gcctgggtgt ggcgctttg
gcgtcctct gccacctgc caagaagcag ggcgccgcca gggccaagat ggagtacaag tgcgcggccc ctccaagga
ggtagtgtg cagcacgtgc ggaccgagc gacgccgag agactc

FIG. 4

SEQ ID NO: 4

GTTTAAACGGGGCCCTCTAGACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGG
GTCATTAGTTCATAGCCCATGATATCATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCC
TGGCTGACCGCCCAACGACCCCCGCCCATTTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAA
TAGGGACTTTTCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATG
CCCAGTNCATGACCTTATGGGACTTTTCTACTTGGCAGACATCTACGTATTAGTCATCGCTATTACCAT
GGTGATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTTCCAAGT
CTCCACCCCATTTGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT
AACAACTCCGCCCCATTGACGCAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAG
CTCTCTGGCTAACTAGAGAACCCCTGCTTACTGGCTTATCGAGATATCTGCAGAATTCATCTGTGCGACT
GCTACCGGCGAGCGCGCAGCGGCAAGAAGTGTCTGGGCTGGGACGGACAGGAGGCTGTGCGCATCG
GCGTCTGTGCCCTCTGCTCCGGCACGGCCCTGTCGCAGTGCCCGCGCTTTCCCGGCGCCTGCACGC
GGCGCGCCTGGGTAACATGCTTGGGGTCTGCTGCTTGGCGCGCTGGCCCTGGCCGGCCTGGGGTTCC
CCGACCCCGCAGAGCCGCGAGCCGGGTGGCAGCCAGTGCCTCGAGCACGACTGCTTCGCGCTCTACCCG
GGCCCCGCGACCTTCTCAATGCCAGTCAGATCTGCGACGGACTGCGGGGCCACCTAATGACAGTGC
CTCCTCGGTGGCTGCCGATGTCAATTTCTTGTACTGAACGGCGACGGCGGCGTGGCCGCCGGCGCCT
CTGGATCGGCCTGCAGCTGCCACCCGGCTGCGGCGACCCCAAGCGCCTCGGGCCCCCTGCGCGGCTTCC
AGTGGGTACGGGAGACAACAACACCAGCTATAGCAGGTGGGCACGGCTCGACCTCAATGGGGCTCC
CCTCTGCGGCCCCGTTGTGCGTCTGTCTCCGCTGCTGAGGCCACTGTGCCAGCGAGCCGATCTGGGA
GGAGCAGCAGTGCGAAGTGAAGGCCGATGGCTTCTCTGCGAGTTCCACTTCCCAGCCACCTGCAGGC
CACTGGCTGTGGAGCCCGCGCCGCGCTGCCGCGCTCTCGATCACCTACGGCACCCCGTTGCGGGCC
CGCGGAGCGGACTTCCAGGCGCTGCCGCTGGGCGAGCTCCGCGCGGTGGCTCCCTCGGCTTACAGCT
AATGTGCACCGCGCCCGCCGAGCGGTCCAGGGGCACTGGGCCAGGGAGGCGCCGGGCGCTTGGGAC
TGCAGCGTGGAGAACGGCGGCTGCGAGCACGCGTGCAATGCGATCCCTGGGGCTCCCCGCTGCCAGTG
CCCAGCCGGCGCCGCCCTGCAGGCAGACGGGCGCTCCTGCACCGCATCCGCGACGCAGTCTGCAACG
ACCTCTGCGAGCACTTCTGCGTTCCCAACCCCGACCAGCCGGGCTCCTACTCGTGATGTGCGAGACC
GGCTACCGGCTGGCGGCCGACCAACACCGGTGCGAGGACGTGGATGACTGCATACTGGAGCCCAGTC
CGTGTCGCGAGCGCTGTGTCAACACACAGGGTGGCTTCGAGTGCCACTGCTACCCTAACTACGACCTG
GTGGACGGCGAGTGTGTGGAGCCCGTGGACCCGTGCTTCAGAGCCAACTGCGAGTACCAGTGCCAGCC
CCTGAACCAAATACTAGCTACCTCTGCGTCTGCGCCGAGGGCTTCGCGCCCCATTCCCCACGAGCCGCACA
GGTGCCAGATGTTTTGCAACCAGACTGCTGCTCCAGCCGACTGCGACCCCAACACCCAGGCTAGCTGT
GAGTGCCATACATCCTGGACGACGCTTTCATCTGCACGGACATCGACAGTGGCGAAAACGG
CGGCTTCTGCTCCGGGGTGTGCCACAACCTCCCCGGTACCTTCGAGTGCACTGCGGGCCCGACTCGGC
CCTTGCCCCGCCACATTGGCACCGACTGTGACTCCGGCAAGGTGGACGGTGGCGACAGCGGCTCTGGCG
AGCCCCCGCCAGCCCGACGCCCGGCTCCACCTTGACTCCTCCGGCCGTGGGGCTCGTGCAATTCGGGC
TTGCTCATAGGCATCTCCATCGCGAGCCTGTGCCTGGTGGTGGCGCTTTTGGCGCTCCTCTGCCACCTG
CGCAAGAAGCAGGGGCGCCGCCAGGGCCAAGATGGAGTACAAGTGCGCGGCCCTTCCAAGGAGGTAG
TGCTGCAGCACGTGCGGACCGAGCGGACCGCGCAGAGACTCTGAGCGGCCTCCGTCCAGGAGCCTGG
CTCCGTCCAGGAGCCTGTGCCTCCTACCCCCAGCTTTGCTACCAAAGCACCTTAGCTGGCATTACAGC
TGGAGAAGACCCTCCCCGCACCCCCCAAGCTGTTTTCTTCTATTCCATGGCTAACTGGCGAGGGGGTG
ATTAGAGGGAGGAGAATGAGCCTCGGCCTCTTCCGTGACGTCACTGGACCACTGGGCAATGATGGCAA
TTTTGTAAACGAAGACACAGACTGCGATTTGTCCAGGTCTCACTACCGGGCGCAGGAGGGTGAGCGT
TATTGGTTCGGCAGCCTTCTGGGCAGACCTTGACCTCGTGGGCTAGGGATGACTAAAATATTTATTTTTT
TTAAGTATTTAGGTTTTTTGTTTGTCTTCTTACCTGTATGTCTCCAGTATCCACTTTGCACAGCT
CTCCGGTCTCTCTCTCTCTACAACTCCCACTTGTCAATGTGACAGGTAAACTATCTTGGTGAATTTTTTT
TTCTAGCCCTCTCACATTTATGAAGCAAGCCCCACTTATTTCCCATTTCTTCTAGTTTTCTCCTCCCAG
GAAGTGGGCCAACTCACCTGAGTCACTTACCTGTGCCTGACCCTACTTCTTTTGTCTTAGCTGTCTG
CTCAGACAGAACCCTACATGAAACAGAAACAAAAACACTAAAAATAAAAAATGGCCATTTGCTTTTTT
ACCAGATTTGCTAATTTATCCTGAAATTTTCAGATTCCCAGAGCAAAATAATTTTAAACAAAGGTTGAG
ATGTAAGAGGTATTAATTTGATGTTGCTGGACTGTCATAGAAATTACACCCAAAGAGGTATTTATCTTT
ACTTTTAAACAGTGAGCCTGAATTTTGTGCTGTTTTGATTTGTACTGAAAAATGGTAATTGTTGCTAA

FIG. 5

TCTTCTTATGCAATTTCTTTTTTGTATTATTACTTATTTTTGACAGTGTTGAAAATGTTTCAGAAGGTT
GCTCTAGATTGAGAGAAGAGACAAACACCTCCCAGGAGACAGTTCAAGAAAGCTTCAAACCTGCATGA
TTCATGCCAATTAGCAATTGACTGTCACTGTTCCCTTGTCAGTGGTAGACCAAAAATAAAACCAGCTCTAC
TGGTCTTGTGGAATTGGGAGCTTGGGAATGGATCCTGGAGGATGCCCAATTAGGGCCTAGCCTTAATC
AGGTCCTCAGAGAATTTCTACCATTTTCAGAGAGGCCTTTTGGAATGTGGCCCCTGAACAAGAATTGGA
AGCTGCCCTGCCCATGGGAGCTGGTTAGAAATGCAGAATCCTAGGCTCCACCCCATCCAGTTCATGAG
AATCTATATTTAACAAGATCTGCAGGGGGTGTGTCTGCTCAGTAATTTGAGGACAACCATTCCAGACT
GCTTCCAATTTTCTGGAATACATGAAATATAGATCAGTTATAAGTAGCAGGCCAAGTCAGGCCCTTATT
TTCAAGAACTGAGGAATTTTCTTTGTGTAGCTTTGCTCTTTGGTAGAAAAGGCTAGGTACACAGCTCT
AGACACTGCCACACAGGGTCTGCAAGGTCTTTGGTTTCAGCTAAGCTAGGAATGAAATCCTGCTTCAGT
GTATGGAAATAAATGTATCATAGAAATGTAACCTTTGTAAGACAAAGGTTTTCTCTTCTATTTTGTAA
ACTCAAAATATTTGTACATAGTTATTTATTTATTGGAGATAATCTAGAACACAGGCAAAATCCTTGCTT
ATGACATCACTTGTACAAAATAAACAAATAACAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAGGTAGCAGTCGACAGATGAATTCACCACACTGGACTAGTGGATCCGAGCTCGGT
ACCAAGCTTAAGTTTAAAC

FIG. 5A

SEQ ID NO 5

TCTAGACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATA
GCCCATGATATCATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCC
AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCA
TTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGC
CAAGTACGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTNCATGAC
CTTATGGGACTTTCCTACTTGGCAGACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTT
TGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTTCCAAGTCTCCACCCCATTG
ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCC
CCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAAC
TAGAGAACCCCTGCTTACTGGCTTATCGAGATATC

FIG. 6

SEQ ID NO 6

GGCAGCGCGCAGCGGCAAGAAAGTGTCTGGGCTGGGACGGACAGGAGAGGCTGTCCGATCGGGCTCC
TGTGCCCCCTCTGCTCCGGCACGGCCCTGTGCGAGTGCCCCGCGCTTTCCCCGGCGCCTGCACGCGGGCGCG
CCTGGGTAACATGCTTGGGGTCCTGGTCCTTGGCGCGCTGGCCCTGGCCGGCCTGGGGTTCCCCGCACC
CGCAGAGCCCGCAGCCGGGTGGCAGCCAGTGCGTTCGAGCAGACTGCTTCGCGCTCTACCCGGGCCCCG
CGACCTTCCTCAATGCCAGTCAGATCTGCGACGGAAGTGGGGGCCACCTAATGACAGTGCGCTCCTCG
GTGGCTGCCGATGTCATTTCTTGTACTGAACGGCGACGGCGGCGTTGGCCGCGGGCGCCTCTGGAT
CGGCCTGCAGCTGCCACCCGGCTGCGGCGACCCCAAGCGCCTCGGGCCCCCTGCGCGGCTTCCAGTGGG
TTACGGGAGACAACAACACCAGCTATAGCAGGTGGGCACGGCTCGACCTCAATGGGGCTCCCCTCTGC
GGCCCGTTGTGCGTCTGTCTCCGCTGCTGAGGCCACTGTGCCAGCGAGCCGATCTGGGAGGAGCA
GCAGTGCGAAGTGAAGGCCGATGGCTTCCTCTGCGAGTTCCACTTCCCAGCCACCTGCAGGCCACTGG
CTGTGGAGCCCCGGCGCCGCGGCTGCCGCCGTCTCGATCACCTACGGCACCCCGTTTCGCGGCCCGCGGA
GCGGACTTCCAGGCGCTGCCGGTGGGCAGCTCCGCCGCGGTGGCTCCCCCTCGGCTTACAGCTAATGTG
CACCAGCCCGCCCGGAGCGGTCCAGGGGCACTGGGCCAGGGAGGCGCCGGGGCGCTTGGGACTGCAGC
GTGGAGAACCGGGCTGCGAGCAGCGCTGCAATGCGATCCCTGGGGCTCCCCGCTGCCAGTGCCGAGC
CGCGCCCGCCCTGCGAGCAGACGGGGCGCTCCTGCACCGCATCCGCGACGCACTCTGCAACGACCTCT
GCGAGCACTTCTGCGTTCCCAACCCCGACCAAGCCGGGCTCCTACTCGTGCATGTGCGAGACCGGCTAC
CGGCTGGCGGGCCGACCAACACCGGTGCGAGGACGTGGATGACTGCATACTGGAGCCCAGTCCGTGTCC
GCAGCGCTGTGTCAACACACAGGGTGGCTTCGAGTGCCACTGCTACCCTAACTACGACCTGGTGGACG
GCGAGTGTGTGGAGCCCGTGGACCCGTGCTTCAGAGCCAACTGCGAGTACCAGTGCCAGCCCGCTGAAC
CAAAGTAGCTACCTCTGCGTCTGCGCCGAGGGCTTCGCGCCCATTCCCACGAGCCGCACAGGTGCCA
GATGTTTTGCAACCAGACTGCCTGTCCAGCCGACTGCGACCCCAACACCCAGGCTAGCTGTGAGTGCC
CTGAAGGCTACATCCTGGACGACGGTTTCATCTGCACGGACATCGACGAGTGCGAAAACGGCGGCTTC
TGCTCCGGGGTGTGCCACAACCTCCCCGGTACCTTCGAGTGCACTCTGCGGGCCCGACTCGGCCCTTGCC
CGCCACATTGGCACCGGACTGTGACTCCGGCAAGGTGGACGGTGGCGACAGCGGCTCTGGCGAGCCCCC
GCCCAGCCCGACGCCCGGCTCCACCTTGACTCCTCCGGCCGTGGGGCTCGTGCACTTCGGGCTTGCTCAT
AGGCATCTCCATCGCGAGCCTGTGCCTGGTGGTGGCGCTTTTGGCGCTCCTCTGCCACCTGCGCAAGAA
GCAGGGCGCCCGCAGGGGCCAAGATGGAGTACAAGTGCAGCGGCCCTTCCAAGGAGGTAGTGCTGCAG
CACGTGCGGACCGAGCGGACGCCGAGAGACTCTGAGCGGCCTCCGTCCAGGAGCCTGGCTCCGTCCA
GGAGCCTGTGCCTCCTACCCCCAGCTTTGCTACCAAAAGCACCTTAGCTGGCATTACAGCTGGAGAAG
ACCCTCCCCGCACCCCCCAAGCTGTTTTCTTCTATTCCATGGCTAACTGGCGAGGGGGTGATTAGAGGG
AGGAGAATGAGCCTCGGCCTCTTCCGTGACGTCACTGGACCACTGGGCAATGATGGCAATTTTGTAAC
GAAGACACAGACTGCGATTTGTCCCAGGTCTCACTACCGGGCGCAGGAGGGTGAGCGTTATTGGTTCG
GCAGCCTTCTGGGCAGACCTTGACCTCGTGGGCTAGGGATGACTAAAATATTTATTTTTTTAAGTATT
TAGGTTTTTTGTTTGTTCCTTTGTTCTTACCTGTATGTCTCCAGTATCCACTTTGCACAGCTCTCCGGTCT
CTCTCTCTACAAACTCCCACTTGTATGTGACAGGTAAACTATCTTGGTGAATTTTTTTTCTAGCC
CTCTCACATTTATGAAGCAAGCCCCCACTTATCCCCATTCTTCTAGTTTTTCTCCTCCAGGAACTGGGC
CAACTACCTGAGTCACCCTACCTGTGCCTGACCCTACTTCTTTTGTCTTTAGCTGTCTGCTCAGACAG
AACCCTACATGAAACAGAAACAAAAACACTAAAAATAAAAAATGGCCATTGCTTTTTTACCAGATTT
GCTAATTTATCCTGAAATTTAGATTCCCAGAGCAAAATAATTTTAAACAAAGGTTGAGATGTAAAG
GTATTAAATTGATGTTGCTGGACTGTATAGAAATTACACCCAAAGAGGTATTTATCTTTACTTTTAAA
CAGTGAGCCTGAATTTTGTGCTGTTTTGATTTGTACTGAAAAATGGTAATTGTTGCTAATCTTCTTATG
CAATTTCTTTTTTTGTTATTATTACTTATTTTTTGACAGTGTTGAAAATGTTTCAAGAGGTTGCTCTAGATT
GAGAGAAGAGACAAACACCTCCCAGGAGACAGTTCAAGAAAGCTTCAAAGTGCATGA
TTCATGCCAATTAGCAATTGACTGTCACTGTTCTTGTCACTGGTAGACCAAAATAAAACCAGCTCTAC
TGGTCTTGTGGAATTGGGAGCTTGGGAATGGATCTGGAGGATGCCCAATTAGGGCCTAGCCCTTAATC
AGGTCTCAGAGAATTTCTACCATTTAGAGAGGCGCTTTTGGGAATGTGGCCCTGAACAAGAAATTGGA
AGCTGCCCTGCCCATGGGAGCTGGTTAGAAATGCAGAATCCTAGGCTCCACCCCATCCAGTTTCATGAG
AATCTATATTTAACAAGATCTGCAGGGGGTGTGTCTGCTCAGTAATTTGAGGACAACCATTCAGACT
GCTTCCAATTTTCTGGAATACATGAAATATAGATCAGTTATAAGTAGCAGGCCAAGTCAGGCCCTTATT
TTCAAGAAACTGAGGAATTTTCTTTGTGTAGCTTTGCTCTTTGGTAGAAAAGGCTAGGTACACAGCTCT
AGACACTGCCACACAGGGTCTGCAAGGTCTTTGGTTTCAAGTAAAGTAAATCCTGCTTCAGT

GTATGGAAATAAATGTATCATAGAAATGTAACTTTTGTAAGACAAAGGTTTTCCTCTTCTATTTTGTA
ACTCAAAATATTTGTACATAGTTATTTATTTATTGGAGATAATCTAGAACACAGGCAAAATCCTTGCTT
ATGACATCACTTGTACAAATAAACAAATAACAATGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAA

FIG. 7A